The Role of Hyponymy and Context Concreteness in Compound Word Processing

Chariton Charitonidis

ORCID iD: 0000-0003-0298-2629

Contact email address: dr.chariton.charitonidis@gmail.com

This is an Accepted Manuscript of an article published by Taylor & Francis in *Studia Neophilologica* on 08 Jun 2024, available online:
https://www.tandfonline.com/doi/full/10.1080/00393274.2024.2336851

The results of this study were first presented at the 25th International Symposium on Theoretical and Applied Linguistics (ISTAL25), held from May 13th to 15th, 2022, in Thessaloniki, Greece. The talk was titled “Context awareness in word recognition: The case of English non-spaced compounds”. The same results were also presented at the 20th International Morphology Meeting (IMM20), held from September 1st to 4th, 2022, in Budapest, Hungary, under the title “Both hyponymy and context concreteness for the second constituent are relevant in the comprehension and production of compound words”.

Abstract

This paper describes the effects of hyponymy and emotion on the comprehension and production of compound words. The research subjects are over 2000 concatenated compounds of English taken from the LADEC database (Gagné et al. 2019). The study builds on the research carried out in Charitonidis (2022), according to which context concreteness for the second constituent was a significant positive predictor of lexical decision and naming times from the English Lexicon Project (ELP) and the British Lexicon Project (BLP). In the present paper, the hyponymy norms from Gagné et al. (2020) were added in the analysis. The results show that both hyponymy and context concreteness for the second constituent are relevant. In addition, all models including both variables have a better fit than nested models omitting one of these variables. There is thus strong evidence that both hyponymy and context concreteness for the second constituent are obligatory parameters in compound word processing.

Keywords: compounding, lexical decision, naming, emotion, concreteness

1 Introduction

Lieber and Štekauer (2009) examined a variety of phonological, syntactic, and morphological criteria to distinguish English compounds from phrases or other sorts of derived words.[1] According to these authors, the strongest hints for establishing a word complex as a compound in English are left-hand stress ('*cart-horse*), inseparability (*[*black ugly bird*] for *blackbird,* a bird species), impossibility of first-stem modification (**a very blackbird*), inability to replace the second stem with a pro-form (*a riding horse… *the carriage ones*), and inflection on the rightmost constituent, i.e. the head (*cart-horse-s*). However, as the same authors argue, none of these hints can be regarded as an absolute criterion for establishing a word complex as a compound (Lieber & Štekauer 2009: 14).
According to Plag (2003: 132), compounding is the most productive word-formation process in English. An inventory of compound types containing two constituent words can be found in Table 1 (Plag 2003: 144). Compounds with more than two constituent words can be broken down into binary left-branching structures, cf. the binary structure [[bathroom towel] designer] for *bathroom towel designer,* etc. (see also Haider 2001).

Table 1: Inventory of compound types in English (Plag 2003: 144)

First constituent	noun (N)	verb (V)	adjective (A)
noun	film society	brainwash	stone-deaf
verb	pickpocket	stir-fry	-
adjective	greenhouse	blindfold	light-green
preposition	afterbirth	-	-

In English compounds with two constituent words, the right-hand constituent is, typically, the categorial and semantic head. Following the definitions in Scalise and Fábregas (2010: 124), the *categorial* head is the unit that defines the lexical category of the whole word (e.g. *whiteboard* (N) < *white* (A) + *board* (N)), and the *semantic* head is the unit that defines the semantic class of the word (e.g. *whiteboard* (object) < *white* (property) + *board* (object)).[2] Regarding semantic headedness, the meaning of the compound is, typically, a hyponym of the meaning of the head (the hypernym), e.g. the meaning of the compound *bedroom* (the more

[1] The descriptions in this section were taken from Charitonidis (2022: section 1). Minor changes were made.
[2] As Jackendoff (2010) reports, 'there are some families of left-headed compounds in English, such as *attorney general*, *mother-in-law*, *blowup*, and *pickpocket*'.

specific term) is a hyponym of the head word *room* (the generic term), etc. (for the notion of 'hyponymy' see Löbner 2013: 205–207).
According to Libben et al. (2020), compounds have a dual nature. They usually contain constituents that are easily identifiable and, at the same time, they are used as unique structures with specific meanings. For the vast majority of compounds, 'if a language user did not previously know the meaning of the whole compound word, it would be very difficult to figure it out on the basis of the meanings of the constituent elements alone' (Libben et al. 2020: 340). However, the activation of both whole-word representations and constituents 'is present whether or not compound words are semantically transparent and whether or not they are written with spaces, without spaces, or with hyphens' (Libben et al. 2020: 349). This dual nature of compound processing is captured by dual- and multiple-route models of word recognition. These models propose that 'the meanings of both complex word and its morphemes can be activated simultaneously', whereby 'the processing preference for either the morphemic or the whole-word route is not categorical and can be biased by the formal properties of the complex word' (Kuperman 2013: 1).[3]
The Large Database of English Compounds (LADEC: Gagné et al. 2019) is the largest existing database of compound words. It contains over 8000 nonspaced ("closed" or "concatenated") compounds (=nouns) selected from various sources. A vast variety of compounds is considered, e.g. noun-noun compounds such as *buttercup, shipyard,* etc., compounds with a second constituent derived from a verbal stem such as *pacemaker, painkiller,* etc. The first non-head constituent refers to a wide range of grammatical categories. Figure 1 contains a brief sample.

Figure 1. LADEC entries: sample

afterlife aircraft ashtray	daydreaming dimwit drawback	pacemaker padlock painkiller
backboard ballplayer buttercup	earthquake egghead eyebrow	shipyard shoelace shotgun
caretaker castaway crossfire	offspring outcasts overdrive	textbook throwback turnaround

In their linear-regression analyses, Gagné et al. (2019) included a wide range of predictor variables, such as letter length (word length in terms of the number of letters), bigram frequency at the morpheme boundary (e.g. the *d* and *l* in padlock), family size (the number of compounds in LADEC that have the constituent of interest), word frequency, probability and association (vector-based) measures, emotional/sentiment norms computed from participant ratings, etc.[4] The compound-based log response times from the English Lexicon Project (ELP: Balota et al. 2007) and the British Lexicon Project (BLP: Keuleers et al. 2012) were used as dependent variables. Gagné et al. (2019) showed that the above-mentioned predictor variables had significant effects on lexical decision and naming times.

[3] Overviews of word recognition models can be found in Schreuder & Baayen (1995), Kuperman (2013), Norris (2013), and Snefjella & Kuperman (2016).
[4] In the literature and in the present paper, the terms 'norms' and 'ratings' are often used interchangeably to refer to the same values. Actually, the output datasets contain *norms* obtained by averaging native speakers' ratings.

2 Previous research

2.1 The role of semantic transparency and hyponymy in English compounding[5]

For determining the role of *semantic transparency* in English compounding, Gagné et al. (2019) asked participants to rate compounds considering how predictable the meaning of the compound is from its parts (*meaning predictability* ratings, compound-based) and how much of the meaning of each of the constituents is retained in the compound (*meaning retention* ratings, constituent-based). The authors found that the distribution of transparencies for the *second* constituent was much more peaked and higher than the distribution of transparencies for the first constituent[6] (M_{C1}: 64.80 [SD: 19.59] vs. M_{C2}: 71.00 [SD: 16.46]. N = 8115).[7] However, the rating for the *first* constituent was more strongly correlated with the rating for the entire compound than was the rating for the second constituent (c1~cmp: $r = 0.75$, $p < .001$ vs. c2~cmp: $r = 0.66$, $p < .001$. N = 429).[8] Most notably, the rating for the first constituent and the rating for the compound predicted all three types of response times, i.e. ELP lexical decision, BLP lexical decision, and ELP naming times.
The peaked and higher distribution of transparencies for the second constituent, along with the first constituent's better association with the compound's meaning predictability, appear to be immediately mapped onto the head operations in English compounds (section 1). The second constituent, i.e. the head, is a unit whose transparency is reinforced categorially and semantically. The first constituent, i.e. the modifier, is the most critical factor in establishing compound reference. Its transparency covaries with the transparency of the whole compound most strongly.[9]
Gagné et al. (2020) obtained human *hyponymy* ratings from 925 native speakers of English for over 2500 English compounds. As for the authors' experimental procedure, 'The hyponymy judgement was presented on a computer screen in the form "Is < compound > a type of < second constituent >"? … Participants responded either yes or no by pressing one of two computer keys' (Gagné et al. 2020: 3). Regarding the authors' analysis, 'The responses were aggregated to obtain the percentage of participants responding yes for each item. The mean judgement was 63.44% (SD = 25.44) with the range from 0.99 to 100%" (Gagné et al. 2020: 4). Gagné et al. (2020) found that hyponymy was a critical positive predictor of the semantic transparency of the first constituent, the second constituent, and the compound as a whole. They also found that hyponymy was a negative, i.e. latency-reducing, predictor of both lexical decision and naming times. The authors conclude that 'hyponymy plays a central role in representation and processing and, importantly, cannot be reduced to any particular existing measure of semantic transparency or association' (Gagné et al. 2020: 13). The fact that hyponymy called for both lexical decision and naming times 'suggests that categorical relationships might play an obligatory role in how people access and use words and morphemes' (Gagné et al. ibid).

---

[5] The descriptions regarding semantic transparency were taken from Charitonidis (2022: section 1.2). Minor changes were made.
[6] This statement indicates that the histogram depicting the data for the second constituent exhibits a taller and narrower peak compared to the histogram representing the data for the first constituent. Furthermore, the distribution for the second constituent contains a higher overall frequency or number of data points than the distribution for the first constituent.
[7] In the parentheses, means and standard deviations are given in percentages. In the present study, the three morphological levels 'compound', 'first constituent', and 'second constituent' are abbreviated as 'cmp', 'c1', and 'c2', respectively.
[8] Steiger's (1980) *z* test showed that this difference was significant, $z = 27.71$, $p < .0001$ (Gagné et al. 2019).
[9] By referring to previous research, Gagné et al. (2019) report that 'the modifier (the first constituent in English) tends to play a larger role in the ease-of-relation selection during the processing of compounds and noun phrases.'

2.2 The role of emotion in English compounding

2.2.1 Obtaining emotion norms for English words[10]

In recent years, there has been a considerable focus on the interface of lexical meaning and emotion (for a review of relevant studies, see Citron et al. 2016 and Yao et al. 2016). Warriner et al. (2013) produced a dataset with norms for English words according to the affective variables 'valence' (positivity) and 'arousal' (excitement, mood-enhancement).[11] Brysbaert et al. (2014) produced a dataset with norms for English words according to the sensorimotor variable 'concreteness'. Henceforth, the word-level norms from these datasets are referred to with the generic term *representation norms.*[12]
In the following, I give the definition of the above-mentioned variables (Kuperman 2013).

> Valence, or emotional positivity, gages the amount of pleasantness or discomfort that a person feels when reading the word, and is measured on a scale from 1 (sad, unhappy) to 9 (happy). Words with extreme average valence ratings are *pedophile* (1.26) and *vacation* (8.53). Arousal assesses the level of excitement that raters associate with the read word, and is measured on a scale from 1 (calm) to 9 (excited). Words with extreme average arousal ratings are *grain* (1.6) and *insanity* (7.79)… Concreteness assesses, on a scale from 1 to 5, how easily the referent of the word can be seen, heard, felt, smelled, or tasted... Words with extreme average concreteness ratings are: *essentialness* (1.04) and *flashlight* (5.00). (Kuperman 2013: 3)

In Snefjella & Kuperman (2016), the application of representation norms to the 7 billion token USENET corpus (Shaoul & Westbury 2013) resulted in valence, arousal, and concreteness norms for word contexts, henceforth referred to as *context norms*. Each context was confined from five content words before to five content words after a target word.[13] Contexts in which fewer than three words matched with representation norms were excluded.[14] Accordingly, 14,853 words were considered that had semantic estimates for both representations and contexts. In Table 2, a sample context for the word *evidence* is given. Blanks indicate the absence of norms for specific words.

Table 2. A sample context for the word evidence (Snefjella & Kuperman 2016: 137)

Word	Valence	Arousal	Concreteness
always			1.71
offer	5.94	3.42	2.23
zero			2.86
factual	5.89	3.05	2.41
logical	6.60	4.11	2.11
evidence	-	-	-
false			2.36
claims	5.15	3.90	
unless			1.54
stupid	2.65	4.68	1.75
unable	2.96	3.76	1.77
Mean	4.87	3.82	2.82

[10] The descriptions in this section were taken from Charitonidis (2022: Introduction). Minor changes were made.
[11] Warriner et al.'s (2013) dataset also included 'dominance' norms. *Dominance* refers to the 'degree of control' exerted by the stimulus word.
[12] In the literature and in the present paper, the terms 'norms' and 'ratings' are used interchangeably to refer to the same values. In fact, the datasets referred to above contain norms obtained through averaging of native speakers' ratings.
[13] In Snefjella & Kuperman (2016), the term 'content words' is equivalent to the term 'non-stopwords'. Stopwords correspond to the default English stopword list of the R tm-package (personal communication).
[14] Also excluded were 493 words whose overall context values 'were more than three standard deviations above or below the mean of the respective variable' (Snefjella & Kuperman 2016: 136).

At the next stage, Snefjella & Kuperman (ibid) averaged all context means across all occurrences of each word in the corpus. The resulting norms refer to three meta-variables, i.e. 'context valence', 'context arousal', and 'context concreteness'.[15] These norms serve as ‘indices of the overall tendency of a word to occur in positive, exciting, or concrete contexts’ (Snefjella & Kuperman 2016: 137).
Snefjella & Kuperman (2016: 139) state that ‘words tend to favour the company of words with similar affective and sensorimotor connotations’. For instance, the noun *athlete* has a representation valence of 6.16 and a context valence of 5.86, i.e. positive values in both cases, or the noun *creatorship* has a representation concreteness of 2.58 and a context concreteness of 1.84, i.e. low values in both cases, etc. In Table 3, the moderate to strong positive correlations of context and word (=representation) ratings refer to this tendency.

Table 3. Correlations of context and word ratings (Snefjella & Kuperman 2016: 139)

Context valence vs. word valence	.58***
Context arousal vs. word arousal	.48***
Context concreteness vs. word concreteness	.72***

***p < .001

2.2.2 Emotion norms vs. lexical decision and naming times (Charitonidis 2022)[16]

Charitonidis (2022) examined the influence of representation and context norms referred to in section 2.2.1 on over 2000 non-spaced (concatenated) English compounds taken from the LADEC database (Gagné et al. 2019). The norms of representation valence and arousal were taken from a modified and expanded version of Warriner et al.'s (2013) dataset, i.e. Kuperman (2020)[17]. As dependent variables were used the ELP lexical decision times, the BLP lexical decision times, and the ELP naming times. Compound length (number of letters) and log compound frequency from the SUBTLEX-US corpus (Brysbaert & New 2009) were used as controls.[18]
The main analysis in Charitonidis (2022) resulted in three “global models” including all significant representation and context predictors from the “local models”, i.e. models referring to a single representation or context variable mapped onto three predictors, each corresponding to one of the three morphological levels, i.e. compound (cmp), first constituent (c1), and second constituent (c2). Table 4 below reproduces the results of the forced-entry regression analysis. Empty cells stand for non-significant predictors in the local models. For the AIC and AIC/N values see section 4.[19]

[15] The full list of norms can be found in the supplementary dataset of Snefjella & Kuperman (2016).
[16] This section draws on various sections in Charitonidis (2022).
[17] This version was also used in Snefjella and Kuperman (2016) and Gagné et al. (2019).
[18] The SUBTLEX-US corpus was based on subtitles from US films and television programs. As Chen et al. (2018) note, ‘a series of recent studies demonstrated that frequency norms derived from subtitles of films and TV programs tended to outperform those from printed texts in accounting for the variance of lexical processing time (and sometimes also accuracy) among native speakers of different languages’ (see Chen et al. 2018: 2 and the references therein).
[19] Table 4 does not include the BLP lexical decision model with the log compound frequencies from the British National Corpus (BNC) as control variable. This model was dissociated from the ELP and BLP lexical decision models with frequencies from the SUBTLEX-US corpus. It should be noted that BNC frequencies were derived from a corpus ‘with a mixture of written and spoken genres’ (Chen et al. 2018: 8).

Table 4. Standardized regression coefficients from global models using representation and context variables to predict English Lexicon Project (ELP) lexical decision (LD) times, British Lexicon Project (BLP) lexical decision times, and ELP naming times (Charitonidis 2022)

Variable	ELP LD	BLP LD	ELP Naming
SUBTLEX-US frequency	-0.332***	-0.436***	-0.411***
Length of compound	0.181***	-0.019	0.270***
Valence (cmp)	-0.140***	-0.200***	
Valence (c1)			
Valence (c2)		0.030	
Context valence (cmp)	-0.016	0.067	-0.040*
Context valence (c1)	-0.059	-0.043	-0.023
Context valence (c2)			-0.023
Arousal (cmp)			
Arousal (c1)	0.002	0.034	
Arousal (c2)	0.041		
Context arousal (cmp)	-0.018	0.004	
Context arousal (c1)	0.019	0.013	0.046*
Context arousal (c2)	0.044		0.058**
Concreteness (cmp)	-0.068	-0.088	
Concreteness (c1)		-0.045	
Concreteness (c2)		-0.030	
Context concreteness (cmp)	-0.007	0.012	
Context concreteness (c1)			
Context concreteness (c2)	0.066*	0.092*	0.058**
N	854	594	2376
AIC	-2841.37	-2231.362	-8417.148
AIC/N	-3.32713	-3.75650	-3.54257

$*p < .05$, $**p < .01$, $***p < .001$.

In the lexical decision models in Table 4 only two variables were successful, i.e. representation valence for the compound and context concreteness for the second constituent. The latter was also successful in the naming model.

The positive coefficients of context concreteness for the second constituent suggest that lower values of this variable facilitate processing whereas higher values of the same variable slow down processing. Charitonidis (2022) assumed that, when the second compound constituent corresponds to a standalone word usually found in contexts of high concreteness, then the strong associative value of this constituent would run counter to accelerating effects of whole-word (compound) representations.

The comparable effects of contextual concreteness on the second constituent in both lexical decision and naming tasks are consistent with *semantic neighborhood* effects, as discussed by Buchanan et al. (2001) and Danguecan & Buchanan (2016). Buchanan et al. (2001) suggest that semantic neighborhoods can predict response times in both lexical decision and naming tasks. The activation of words extends to those that are highly interconnected. These highly interconnected words do not necessarily share similar features. A significant presence of related concrete words is typically observed in contexts of high concreteness. Alternatively, a substantial presence of related concrete words may lead to latency-inducing *competition effects*, as proposed by Danguecan & Buchanan (2016). These assumptions are also compatible with Paivio's (1978; 2007) theory. According to this theory, concrete words have a strong associative value because they are doubly, i.e. linguistically and perceptually, encoded.[20]

Speeded naming is a shallow task that, typically, does not involve word semantics and can be accomplished purely on a formal basis (see Kuperman 2013 and the references therein). Accordingly, in the global ELP naming model in Table 4 no representation (i.e. word-level) predictors show up. In contrast, four context predictors were successful, i.e. context valence for

[20] The potential inhibitory function of the second compound constituent, as outlined in this paragraph, was first explored in Charitonidis (2022: sections 5.1.2 and 7).

the compound, context arousal for the first and second constituent, and context concreteness for the second constituent.

3 The present study

As already presented, (a) Gagné et al. (2020) reported accelerating effects of hyponymy on both lexical decision and naming times (section 2.1), and (b), Charitonidis (2022) reported inhibitory effects of context concreteness for the second compound constituent on both lexical decision and naming times (section 2.2.2). Charitonidis (2022: 18) argues that in order to uphold the hyponymy relation that calls for the referential operations of the first constituent, it is imperative for the second constituent to maintain a certain level of contextual abstraction. A second constituent that typically emerges in contexts of high concreteness as a free morpheme would impede the referential workings of the first constituent. Based on these premises, the present study seeks to find whether hyponymy and context concreteness for the second constituent are equally relevant across different tasks.

In particular, if our analysis shows that both variables predict lexical decision and naming times jointly while improving the goodness-of-fit of models, then there will be strong evidence that both variables are obligatory parameters in compound processing, potentially redefining the notion of categorial and semantic head in English compounds (section 1).

The research questions are:

1. What are the relationships among the norms of hyponymy (Gagné et al. 2020), transparency (Gagné et al. 2019; 2020), representation-based emotion (Brysbaert et al. 2014; Kuperman 2020), and context-based emotion (Snefjella & Kuperman 2016)?
2. Were the significant effects of context predictors in the global naming model in Charitonidis (2022) due to the omission of hyponymy as a stronger word-level (representation) parameter?
3. Does the inclusion of both hyponymy and context concreteness for the second constituent in response time models consistently enhance model fit across different measures?

The motivation behind these three research questions is primarily theoretical. Investigating them will deepen our comprehension of the fundamental nature of two parameters: hyponymy and context concreteness for the second constituent. There is a likelihood that noun phrases, multiword expressions, and other linguistic constructs share these parameters with compounds. Our methods will be presented in section 4. In section 5.1, the Pearson correlations among the norms of hyponymy, transparency, and emotion will be presented. In section 5.2, a set of nine linear-regression models will be built, using hyponymy and context concreteness for the second constituent in different combinations. Two different measures of goodness of fit will be applied to favour the best models. The results will be discussed in section 6.

4 Methods

The research subjects were over 2000 English concatenated compounds taken from the LADEC database (Gagné et al. 2019).[21] Our methods were based on the linear regression analyses in Gagné et al. (2019; 2020) and Charitonidis (2022). In particular, the forced-entry regression method was used, according to which all predictors were entered simultaneously to model lexical decision times from ELP (Balota et al. 2007) and BLP (Keuleers et al. 2012), and naming times from ELP.

The model construction was based on the global ELP and BLP models in Charitonidis (2022) presented in section 2.2.2. After excluding non-significant predictors, the tests for these models were rerunned to obtain baseline models (see column (1) in Tables 6–8, section 5.2). The

[21] On the website Researchgate.net can be found (a) the full set of compounds with the corresponding input values, and (b) the descriptions and sources for all variables.

variables of ‘hyponymy’ and ‘context concreteness for the second constituent’ were used categorically in different combinations.[22] As in Charitonidis (2022), the SUBTLEX-US log frequency for the compounds and compound length (in characters) were used as control variables. Constituent frequencies were excluded to maintain consistency with the models used in the studies by Gagné et al. (2019) and Charitonidis (2022), which primarily focused on compound-level variables as controls.

For assessing model fit, AIC values were calculated. The AIC equation used was $AIC = -2\ln L + 2k$, in which $\ln L$ refers to the maximized/full log-likelihood of the model and k refers to the number of parameters including the constant. The lower (=the more negative) the AIC value, the better the fit of the model. To facilitate model comparison, a scaled AIC value for each model was computed by dividing AIC with sample size (AIC/N).[23] To validate the AIC/N values, additional Wald tests were run.

5 Results

5.1 Correlations

The purpose of this section is to integrate hyponymy into the broader context of transparency and emotion variables. The primary aim is to investigate whether and how different variables target the right-hand semantic and categorial head within English compounds.

Table 5 below contains the Pearson correlations among norms of hyponymy (Gagné et al. 2020), transparency (Gagné et al. 2019; 2020), representation-based emotion (Brysbaert et al. 2014; Kuperman 2020), and context-based emotion (Snefjella & Kuperman 2016). The correlations referred to a sample of 856 items (listwise deletion).[24] Henceforth, I will discuss the significant correlations alone. Correlations below .3 will be regarded as weak, correlations from .3 to .7 as moderate, and correlations higher than .7 as strong.

A. The positive correlations between hyponymy judgment and (a) predictability (compound transparency), i.e. r = .59***, (b) meaning retention of c1, i.e. r = .21***, and (c) meaning retention of c2, i.e. r = .83***, were almost identical to that in Gagné et al. (2020) (r = .56***, .25***, .81***, respectively. N=2753).

B. The positive correlations between the transparency measures themselves, i.e. r = .77*** (cmp-c1), r = .68*** (cmp-c2), and r = .23*** (c1-c2), were similar to that in Gagné et al. (2020; r = .80***, .65***, .31***, respectively. N=2753). As can be seen, the correlation between c1 and c2 was weak in our tests in contrast to the corresponding moderate correlation in Gagné et al. (2020).

C. Regarding the relation of hyponymy judgement and emotion, there was a single moderate correlation between the former and representation concreteness for the compound (r = .47***). This correlation was followed by a weak correlation between hyponymy judgement and context concreteness for the compound (r = .25***).

---

[22] In other words, hyponymy and context concreteness for the second constituent were utilized with both positive and negative features (+/–).

[23] The Akaike Information Criterion (AIC) encounters challenges when dealing with different sample sizes (Burnham & Anderson 2002). All information criteria rely on the likelihood function, which is influenced by sample size. As sample size increases, likelihood decreases, leading to higher (=inferior) information criterion values. To mitigate this issue, this study employs scaled AIC, dividing AIC by the sample size. While not universally embraced, this method is commonly utilized across diverse applications and aids in adjusting for discrepancies in sample sizes when evaluating models (see, for instance, Hastie et al. 2009: 230-231; for further details, see Charitonidis 2022).

[24] In listwise deletion, only data points with complete information for all variables are used. In the present analysis, listwise deletion resulted in a substantial number of cases (856 items out of the original sample) and is a viable option. It should be noted that Gagné et al. (2019) also used listwise deletion in their correlation analysis.

D. Regarding the relation of transparency norms and emotion, there was a single moderate correlation between the meaning retention rating for the second constituent and representation concreteness for the compound (r = .37***), followed by a weak correlation between the meaning retention rating for the second constituent and context concreteness for the compound (r = .21***).[25]

The patterns in C. and D. suggest that (a) both hyponymy and semantic transparency are dissociated from the affective variables 'valence' and 'arousal', and (b) both hyponymy and the meaning retention rating for the second constituent interact positively and moderately with representation concreteness for the compound. The finding in (b) considers the compound as a whole unit with its own specific meaning and potentially unique level of concreteness.

In section 2.1 it was reported that in Gagné et al. (2020) hyponymy was a critical positive predictor of the semantic transparency of the first constituent, the second constituent, and the compound as a whole. Additionally, the correlation between hyponymy judgment and the meaning retention rating of the second constituent, as mentioned earlier, was notably robust (r = .83***), prompting concerns regarding potential multicollinearity. These patterns suggest that while the semantic transparency of the second constituent may target the semantic and categorial head, it essentially overlaps with the concept of hyponymy. Nevertheless, as noted by Gagné et al. (2020: 12), "hyponymy is a more fundamental relation than semantic transparency in the sense that it applies even when there is no compounding involved." Therefore, to alleviate redundancy concerns, semantic transparency was excluded from the analysis.

[25] Regarding the emotion variables themselves, most of the strongest correlations were between representation and context norms (see also Charitonidis 2022). In particular, all representation norms correlated moderately and positively with context norms at the same morphological level (cmp, c1, c2). The correlation between representation and context concreteness was moderate, i.e. r = .56*** (In Snefjella & Kuperman 2016, the correlation between the same variables was strong, i.e. r = .72***, see Table 3 in section 2.2.1). Context valence and context arousal correlated moderately and negatively at the same morphological level, whereby the correlations between representation valence and representation arousal were weak and negative, again at the same morphological level.

Table 5. Pearson correlations among norms of hyponymy (Gagné et al. 2020), transparency (Gagné et al. 2019; 2020), representation-based emotion (Brysbaert et al. 2014; Kuperman 2020), and context-based emotion (Snefjella & Kuperman 2016)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Hyponymy judgement	—																					
2. Predictability	.59***	—																				
3. Meaning retention (c1)	.21***	.77***	—																			
4. Meaning retention (c2)	.83***	.68***	.23***	—																		
5. Valence (cmp)	.11**	.15***	.11**	.10**	—																	
6. Valence (c1)	.11**	.11**	.09**	.07	.50***	—																
7. Valence (c2)	.10**	.02	-.01	.08*	.28***	.09*	—															
8. Arousal (cmp)	-.15***	-.01	.01	-.12***	-.21***	-.14***	-.09**	—														
9. Arousal (c1)	-.06	.02	.05	-.03	-.12***	-.22***	.01	.44***	—													
10. Arousal (c2)	.04	.03	-.01	.03	-.08*	-.03	-.12***	.28***	.10**	—												
11. Concreteness (cmp)	.47***	.28***	.12***	.37***	.13***	-.01	.00	-.17***	-.06	-.03	—											
12. Concreteness (c1)	.13***	.07*	.00	.08*	-.06	-.02	-.07*	-.10**	-.13***	-.01	.31***	—										
13. Concreteness (c2)	.17***	-.03	-.13***	.13***	-.09**	-.12***	.03	-.10**	-.04	-.03	.36***	.09**	—									
14. Context valence (cmp)	.08*	.03	.00	.06	.56***	.29***	.27***	-.20***	-.14***	-.13***	.10**	.00	.02	—								
15. Context valence (c1)	.11**	.08*	.05	.07*	.34***	.59***	.10**	-.17***	-.26***	-.02	.02	-.05	.01	.26***	—							
16. Context valence (c2)	.09**	.05	.02	.05	.22***	.11**	.42***	-.08*	-.02	-.12***	.00	.03	-.10**	.20***	.10**	—						
17. Context arousal (cmp)	-.09**	-.02	-.00	-.08*	-.20***	-.14***	-.05	.51***	.32***	.26***	-.17***	-.07*	-.03	-.36***	-.21***	-.10**	—					
18. Context arousal (c1)	-.02	-.06	-.08*	-.02	-.13***	-.17***	.01	.31***	.51***	.09*	-.13***	-.14***	-.08*	-.18***	-.46***	-.06	.32***	—				
19. Context arousal (c2)	.01	.01	-.03	.06	-.04	-.07	.03	.20***	.12***	.50***	.01	-.11**	.03	-.01	-.07*	-.34***	.25***	.11**	—			
20. Context concreteness (cmp)	.25***	.10**	-.03	.21***	.21***	.08*	.05	-.14***	-.01	-.03	.56***	.21***	.24***	.24***	.08*	-.00	-.16***	-.06	-.02	—		
21. Context concreteness (c1)	.10**	.01	-.07	.06	.08*	.11**	.00	-.14***	-.16***	-.03	.28***	.56***	.16***	.09*	.21***	.03	-.11**	-.22***	-.09**	.29***	—	
22. Context concreteness (c2)	.14***	-.03	-.10**	.07	-.04	-.07*	.08*	-.07	-.04	-.05	.29***	.13***	.56***	.03	.06	.06	-.02	-.07*	-.12***	.26***	.21***	—

N = 856. *p < .05, **p < .01, ***p < .001 (two-tailed)

5.2 Hyponymy vs. emotion: Regressions

5.2.1 English Lexicon Project (ELP) lexical decision models

In Table 6 below, the standardized regression coefficients and standard errors (in parentheses) from the ELP lexical decision models are given, including representation valence for the compound, hyponymy, and context concreteness for the second constituent as the main variables of interest. As can be seen, representation valence for the compound and hyponymy were associated to faster response times, whereas context concreteness for the second constituent was associated to longer response times. Regarding goodness of fit, Model 2 including both hyponymy and context concreteness for the second constituent was favoured, AIC/N = -3.36281. Model 3 including hyponymy alone yielded an AIC/N value of -3.34845. This value was better than that for the baseline Model 1 including only context concreteness for the second constituent, AIC/N = -3.30375.[26]

Table 6. Standardized coefficients with standard errors from English Lexicon Project (ELP) lexical decision models using hyponymy and context concreteness for the second constituent in different combinations

	(1)	**(2)**	**(3)**
SUBTLEX-US frequency	-0.296*** (0.003)	-0.291*** (0.003)	-0.302*** (0.003)
Length of compound	0.223*** (0.001)	0.177*** (0.001)	0.181*** (0.001)
Valence (cmp)	-0.180*** (0.001)	-0.184*** (0.001)	-0.174*** (0.001)
Context concreteness (c2)	0.060* (0.007)	0.103*** (0.008)	
Hyponymy judgment		-0.061* (0.000)	-0.032 (0.000)
N	1262	1058	1105
AIC	-4169.334	-3557.85	-3700.038
AIC/N	-3.30375	-3.36281	-3.34845

*$p < .05$, **$p < .01$, ***$p < .001$

5.2.2 British Lexicon Project (BLP) lexical decision models

In Table 7 below, the standardized regression coefficients and standard errors (in parentheses) from the BLP lexical decision models are given. The models included the same predictor variables as the ELP lexical decision models in section 5.2.1. In a similar way as in the ELP lexical decision models, (a) representation valence for the compound and hyponymy predicted faster response times while context concreteness for the second constituent predicted longer response times, (b) Model 2 including both hyponymy and context concreteness for the second constituent had the best fit, AIC/N = -3.79552, and (c) Model 3 including hyponymy alone was better than Model 1 including only context concreteness for the second constituent, see the AIC/N values -3.76718 and -3.76618, respectively.[27]

[26] It should be noted that the coefficient for hyponymy in Model 3 did not reach statistical significance, $\beta = -0.032$, $p = 0.246$.
[27] The coefficient for hyponymy in Model 3 reached statistical significance, $\beta = -0.073^*$, $p = 0.026$, in contrast to the corresponding ELP lexical decision model (see section 5.2.1).

Table 7. Standardized coefficients with standard errors from British Lexicon Project (BLP) lexical decision models using hyponymy and context concreteness for the second constituent in different combinations

	(1)	(2)	(3)
SUBTLEX-US frequency	-0.401*** (0.002)	-0.373*** (0.003)	-0.381*** (0.003)
Length of compound	-0.038 (0.001)	-0.055 (0.001)	-0.051 (0.001)
Valence (cmp)	-0.207*** (0.001)	-0.159*** (0.001)	-0.155*** (0.001)
Context concreteness (c2)	0.062* (0.006)	0.124*** (0.007)	
Hyponymy judgment		-0.116** (0.000)	-0.073* (0.000)
N	1041	765	793
AIC	-3920.592	-2903.574	-2987.37
AIC/N	-3.76618	-3.79552	-3.76718

*$p < .05$, **$p < .01$, ***$p < .001$

5.2.3 English Lexicon Project (ELP) naming models

In English compounding, the semantic relation of hyponymy can be mapped onto the formal head-nonhead configuration (see section 1). At the same time, speeded naming is a formal task that largely circumvents representation (word-level) semantics. It was thus expected that in naming models the interplay of hyponymy with context concreteness for the second constituent would still be relevant, as in the ELP and BLP lexical decision models. As will become apparent, our expectations were met.

In Table 8 below, the standardized regression coefficients and standard errors (in parentheses) from the ELP naming models are given. As can be seen, hyponymy and context valence for the compound were mapped onto significant negative predictors, whereas context arousal for both constituents and context concreteness for the second constituent were mapped onto significant positive predictors. The inclusion of both hyponymy and context concreteness for the second constituent in Model 2 resulted in a better AIC/N value in comparison to Model 3 in which hyponymy was solely included, see the values -3.58686 and -3.58379, respectively. Consistent with the observations in sections 5.2.1 and 5.2.2, the AIC/N value for Model 3 was better than that for the baseline Model 1 (-3.54304) that solely included context concreteness for the second constituent.

Table 8. Standardized coefficients with standard errors from English Lexicon Project (ELP) naming models using hyponymy and context concreteness for the second constituent in different combinations

	(1)	(2)	(3)
SUBTLEX-US frequency	-0.413*** (0.001)	-0.421*** (0.001)	-0.426*** (0.001)
Length of compound	0.268*** (0.001)	0.276*** (0.001)	0.276*** (0.001)
Context valence (cmp)	-0.048** (0.004)	-0.042* (0.004)	-0.040* (0.004)
Context arousal (c1)	0.057** (0.006)	0.037* (0.007)	0.037* (0.007)
Context arousal (c2)	0.068*** (0.006)	0.045* (0.007)	0.035 (0.007)
Context concreteness (c2)	0.057** (0.005)	0.055** (0.005)	
Hyponymy judgement		-0.054** (0.000)	-0.047* (0.000)
N	2376	2062	2062
AIC	-8418.27	-7396.108	-7389.784
AIC/N	-3.54304	-3.58686	-3.58379

*$p < .05$, **$p < .01$, ***$p < .001$

5.2.4 Wald tests

Hitherto, the focus was placed on the AIC/N measure to assess model fit. In the lexical decision and naming models, including both hyponymy and context concreteness for the second constituent yielded the most negative (=better) AIC/N value. It was thus expected that *nested,* i.e. reduced, models with only one of these two variables, will always be linked to a significant reduction of the coefficient of determination (R^2).[28] The 'Wald tests' detecting this change have specific advantages in hierarchical regression analyses (Gagné et al. 2020: 4–5).
As illustrated in Table 9 below, models excluding hyponymy ('Model 2') and models excluding context concreteness for the second constituent ('Model 3') were compared against the full models ('Model 1'). It is evident that omitting either variable resulted in inferior model performance.

Table 9. Wald tests for nested models omitting hyponymy ('Model 2') or context concreteness for the second constituent ('Model 3') to predict English Lexicon Project (ELP) lexical decision (LD) times, British Lexicon Project (BLP) lexical decision times, and ELP naming times

Model	R^2 square	F change	df1	df2	p
ELP LD					
1	.184[a]	47.506	5	1052	.000
2	-.004	4.562	1	1052	.033
3	-.010	13.175	1	1052	.000
BLP LD					
1	.211[a]	40.479	5	759	.000
2	-.013	12.091	1	759	.001
3	-.015	14.007	1	759	.000
ELP Naming					
1	.288[b]	118.711	7	2054	.000
2	-.003	8.286	1	2054	.004
3	-.003	8.309	1	2054	.004

a. Predictors: (Constant), hyponymy judgement, length of compound, SUBTLEX-US frequency, representation valence (cmp), context concreteness (c2)
b. Predictors: (Constant), hyponymy judgement, length of compound, SUBTLEX-US frequency, context valence (cmp), context arousal (c1), context arousal (c2), context concreteness (c2)

To conclude, the AIC/N measures in sections 5.2.1–5.2.3 and the R^2-change measures in this section favour the same models.

6 Discussion

With reference to the research questions in section 3, the results of this study are:
1. The positive moderate-to-strong correlations between norms of hyponymy and semantic transparency, as well as between the transparency measures themselves confirmed Gagné et al.'s (2020) analyses. In addition, both hyponymy and semantic transparency were dissociated from the affective variables 'valence' and 'arousal'. In contrast, both hyponymy and semantic transparency for the second constituent interacted positively and moderately with representation concreteness for the compound (section 5.1).
2. There was strong evidence that the effects of context predictors in the global naming model in Charitonidis (2022) were *not* due to the omission of a stronger representation parameter. In particular, even after the addition of a hyponymy predictor in this model, all context predictors remained significant (section 5.2.3).

[28] The *coefficient of determination (R^2)* is an effect-size estimation referring to 'the percentage of variance in one variable that is predicted or explained by the other' (Ozer 1985: 307).

3. Including both hyponymy and context concreteness for the second constituent to predict lexical decision and naming times always led to better models (section 5.2.4). As already pointed out in section 1, in standard linguistic descriptions it is not entirely clear how English compounds are defined as linguistic objects. However, the presence of a category-defining second constituent together with the semantic relation of hyponymy provide a general framework for a principled definition.
In the present analysis, the author employed hyponymy alongside word-level (“representation”) and context variables of emotion to predict lexical decision and naming times. The formal parameters ‘compound length’ (in characters) and ‘log compound frequency’ were used as control variables. In the lexical decision models were successful (a) latency-reducing representation valence for the compound, (b) latency reducing hyponymy, and (c) latency-inducing context concreteness for the second constituent. In the naming models, representation valence for the compound was not successful as opposed to context variables. Most notably, however, both lexical decision and naming models including both hyponymy and context concreteness for the second constituent had a better fit than nested models omitting either of these variables. There is thus strong evidence that hyponymy and context concreteness for the second constituent are obligatory parameters in compound processing.
Similar to the analysis in Charitonidis (2022), the present paper confirmed the dual nature of compound processing by including both semantic and formal parameters in response time models. Future research should examine the relevance of hyponymy and context concreteness for the second constituent in juxtaposition to other variables that in various studies had significant effects on both lexical decision and naming, such as ‘sentiment probability’ (Gagné et al. 2019), ‘imageability’ (Balota et al. 2004: 298, 312), ‘age of acquisition’ (Cortese & Khanna 2007), etc.
In conclusion, the findings of this study have significant implications for how compound words are properly defined or understood. In particular, the present research supports a definition of compound words that emphasizes the inseparability of formal, semantic, and context parameters. Compounds are not simply the sum of their parts, but rather unique entities with integrated semantic representations from both constituents.

Disclosure statement
No potential conflict of interest was reported by the author.

Notes on contributor
Chariton Charitonidis is an independent researcher. His research interests are in the areas of word formation, multiword expressions, lexical semantics, word recognition, and emotion.

ORCID
Chariton Charitonidis ORCID iD 0000-0003-0298-2629

REFERENCES

Balota, David A., Michael J. Cortese, Susan D. Sergent-Marshall, Daniel H. Spieler & Melvin J. Yap. 2004. Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General* 133(2). 283-316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, David A., Melvin J. Yap, Michael J. Cortese, Keith I. Hutchison, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson & Rebecca Treiman. 2007. The English Lexicon Project. *Behavior Research Methods* 39. 445-459. https://doi.org/10.3758/ BF03193014

Brysbaert, Marc & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41. 977-990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, Marc, Amy Beth Warriner & Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46. 904-911. https://doi.org/10.3758/s13428-013-0403-5

Buchanan, Lori, Chris Westbury & Curt Burgess. 2001. Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review* 8(3). 531-544. https://doi.org/10.3758/BF03196189

Burnham, Kenneth P. & David R. Anderson. 2002. *Model selection and multimodel inference: A practical information-theoretic approach,* 2nd edn. New York: Springer. https://dx.doi.org/10.1007/b97636

Charitonidis, Chariton. 2022. Context concreteness for the second constituent slows down compound-word processing. *Lexis* 20. https://doi.org/10.4000/lexis.6769

Chen, Xiaocong, Yanping Dong & Xiufen Yu. 2018. On the predictive validity of various corpus-based frequency norms in L2 English lexical processing. *Behavior Research Methods* 50. 1-25. https://doi.org/10.3758/s13428-017-1001-8

Citron, Francesca M. M., Cristina Cacciari, Michael Kucharski, Luna Beck , Markus Conrad & Arthur M. Jacobs. 2016. When emotions are expressed figuratively: Psycholinguistic and affective norms of 619 idioms for German (PANIG). *Behavior Research Methods* 48. 91-111. https://doi.org/10.3758/s13428-015-0581-4

Cortese, Michael J. & Maya M. Khanna. 2007. Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology* 60(8). 1072-1082. https://doi.org/10.1080/17470210701315467

Danguecan, Ashley N. & Lori Buchanan. 2016. Semantic neighborhood effects for abstract versus concrete words. *Frontiers in Psychology* 7 (Article 1034). https://doi.org/10.3389/fpsyg.2016.01034

Gagné, Christina L., Thomas L. Spalding & Daniel Schmidtke. 2019. LADEC: The Large Database of English Compounds. *Behavior Research Methods* 51. 2152-2179. https://doi.org/10.3758/s13428-019-01282-6

Gagné, Christina L., Thomas L. Spalding, Patricia Spicer, Dixie Wong, Beatriz Rubio & Karen Perez Cruz. 2020. Is buttercup a kind of cup? Hyponymy and semantic transparency in compound words. *Journal of Memory and Language* 113. https://doi.org/10.1016/j.jml.2020.104110.

Haider, Hubert. 2001. Why are there no complex head-initial compounds? In Chris Schaner-Wolles, John Rennison & Friedrich Neubart (eds.), *Naturally!*, 165-174. Torino: Rosenberg & Sellier. https://doi.org/10.1017/S1470542703250261

Hastie, Trevor, Robert Tibshirani & Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edn. New York: Springer.

Jackendoff, Ray. 2010. The ecology of English noun-noun compounds. In Ray Jackendoff (ed.), *Meaning and the lexicon: The parallel architecture 1975-2010,* 413-451. Oxford: Oxford University Press.

Keuleers, Emmanuel, Paula Lacey, Kathleen Rastle & Marc Brysbaert. 2012. The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods* 44. 287-304. https://doi.org/10.3758/s13428-011-0118-4

Kuperman, Victor. 2013. Accentuate the positive: Diagnostics of semantic access in English compounds. *Frontiers in Language Sciences* 4 (Article 203). https://doi.org/10.3389/fpsyg.2013.00203

Kuperman, Victor. 2020. *Modified and expanded norms for valence, arousal and dominance* [Data set]. https://osf.io/zj3u8 (last accessed on 19 April 2023).

Libben, Gary, Christina L. Gagné & Wolfgang U. Dressler. 2020. The representation and processing of compounds words. In Vitto Pirrelli, Ingo Plag & Wolfgang U. Dressler (eds.), *Word knowledge and word usage,* 336-352. Berlin & Boston: de Gruyter. https://doi.org/10.1515/9783110440577-009

Lieber, Rochelle & Pavol Štekauer. 2009. Introduction: Status and definition of compounding. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding,* 3-18. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199695720.001.0001

Löbner, Sebastian. 2013. *Understanding semantics,* 2nd edn. London: Routledge. https://doi.org/10.4324/9780203528334

Norris, Dennis. 2013. Models of visual word recognition. *Trends in Cognitive Sciences* 17(10). 517-524. https://doi.org/10.1016/j.tics.2013.08.003

Ozer, Daniel J. 1985. Correlation and the coefficient of determination. *Psychological Bulletin* 97. 307-315. https://doi.org/10.1037/0033-2909.97.2.307

Paivio, Allan. 1978. The relationship between verbal and perceptual codes. In Edward C. Carterette & Morton P. Friedman (eds.), *Perceptual coding,* 375-397. New York & London: Academic Press. https://doi.org/10.1016/C2013-0-10468-0

Paivio, Allan. 2007. *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Plag, Ingo. 2003. *Word-formation in English.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841323

Scalise, Sergio & Antonio Fábregas. 2010. The head in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding,* 109-125. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.311.10sca

Schreuder, Robert & R. Harald Baayen. 1995. Modeling morphological processing. In Laurie B. Feldman (ed.), *Morphological aspects of language processing,* 131-154. Hillsdale NJ: Lawrence Erlbaum Associates.

Shaoul, Cyrus & Chris Westbury. 2013. *A reduced redundancy usenet corpus (2005-2011)* [Data set]. Edmonton, AB: University of Alberta. https://www.psych.ualberta.ca/~westburylab/downloads/usenetcorpus.download.html (last accessed on 19 April 2023).

Snefjella, Bryor & Victor Kuperman. 2016. It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition* 156. 135-146. https://doi.org/10.1016/j.cognition.2016.07.010

Steiger, James H. 1980. Tests for comparing elements of a correlation matrix. *Psychological Bulletin* 87. 245-251. https://doi.org/10.1037/0033-2909.87.2.245

Warriner, Amy Beth, Victor Kuperman & Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45(4). 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

Yao, Zhao, Jia Wu, Yanyan Zhang & Zhenhong Wang. 2016. Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods* 49. 1374-1385. https://doi.org/10.3758/s13428-016-0793-2

Large Language Models

GPT-3.5 (https://chat.openai.com) and Google Gemini (https://gemini.google.com) were selectively utilized as auxiliary proofreading and editing aids (accessed March 14, 2024). Subsequently, the responses generated by both AI tools underwent additional proofreading and editing by the author.